ISSN 1561-5359. Штучний інтелект, 2018, № 4

УДК 004.032.26

Е.Е. Марушко, А.А. Дудкин
Объединенный институт проблем информатики Национальной академии наук Беларуси, Беларусь
ул. Сурганова, 6, г. Минск, 22001

ИСПОЛЬЗОВАНИЕ АНСАМБЛЯ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОБРАБОТКИ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

Y. Marushko, A. Doudkin
United Institute of Informatics Problems of National Academy of Sciences of Belarus, Belarus
6, Surganova St., Minsk 220012

USING AN ENSEMBLE OF CONVOLUTIONAL NEURAL NETWORKS TO PROCESS REMOTE SENSING DATA

В работе предлагается методика на основе ансамблей моделей для идентификации объектов дистанционного зондирования Земли. В качестве модели предлагается использовать многослойную сверточную нейронную сеть. Определение оптимальных гиперпараметров модели предлагается выполнять методами сеточного или случайного поиска с использованием k-кратной перекрёстной проверки. Показан способ повышения точности идентификации объектов с использованием ансамбля нейронных сетей. Представлены результаты проведения эксперимента с использованием данных дистанционного зондирования Земли. Решалась задача идентификации объектов двух классов, в качестве тестовых данных использовались изображения, полученные с помощью радара с синтезированной апертурой.

Ключевые слова: сверточная нейронная сеть, ансамбль нейронных сетей, дистанционное зондирование Земли, классификация

The paper proposes a technique based on an ensemble of models for identifying objects of remote sensing of the Earth. As a model, it is proposed to use a multilayer convolutional neural network. It is proposed to perform the determination of the optimal hyperparameters of the model using the grid or random search methods using k-fold cross-validation. A method for improving the accuracy of identifying objects using an ensemble of neural networks is shown. The results of the experiment using remote sensing data of the Earth are presented. The task of identifying objects of two classes was solved, images obtained using a synthetic aperture radar were used as test data.

Keywords: convolutional neural network, neural network ensemble, remote sensing of the Earth, classification

Введение

Дистанционное зондирование Земли (ДЗЗ) - наблюдение поверхности Земли наземными, авиационными и космическими средствами, оснащёнными различными видами съемочной аппаратуры. Рабочий диапазон длин волн, принимаемых съёмочной аппаратурой, составляет от долей микрометра (видимое оптическое излучение) до метров (радиоволны). Методы зондирования могут быть пассивные, то есть использующие естественное отраженное или вторичное тепловое излучение объектов на поверхности Земли, обусловленное солнечной активностью, и активные, использующие вынужденное излучение объектов, инициированное искусственным источником направленного действия. Данные ДЗЗ, полученные с космического аппарата (КА), характеризуются большой степенью зависимости от прозрачности атмосферы. Поэтому на КА используется многоканальное оборудование пассивного и активного типов, регистрирующее электромагнитное излучение в различных диапазонах.

Постановка проблемы

При ДЗЗ применяется обработка цифровых данных, представленных в виде двухмерного изображения, которое в одном спектральном диапазоне можно представить в виде матрицы (двухмерного массива) чисел $I(i, j)$, каждое из которых представляет интенсивность излучения, принятого датчиком от элемента по-

верхности Земли, которому соответствует один пиксель изображения.

Изображение состоит из $n \times m$ пикселей, каждый пиксель имеет координаты (i, j) – номер строки и номер колонки. Если изображение получено в нескольких диапазонах электромагнитного спектра, то его представляет трехмерная решетка, состоящая из чисел I (i, j, k), где k – номер спектрального канала. Таким образом, информация, получаемая при ДЗЗ, представляет собой данные с пространственными связями между признаками I (i, j). Для подобного рода данных в последние годы наиболее успешно применялись глубокие сверточные нейронные сети [5].

Сверточные нейронные сети

Архитектура сверточных нейронных сетей (СНС) была предложена Яном Лекуном и нацелена на эффективное распознавание изображений. Свое название архитектура сети получила из-за наличия операции свертки, суть которой заключается в том, что каждый фрагмент изображения умножается на матрицу (ядро) свертки поэлементно, а результат суммируется и записывается в аналогичную позицию изображения [7].

СНС обычно представляет собой чередование сверточных слоев, субдискретизирующих слоев и полносвязных слоев. Все три вида слоев могут чередоваться в произвольном порядке.

Полносвязный слой (dense) - слой, в котором каждый нейрон соединен со всеми нейронами на предыдущем уровне, причем каждая связь имеет свой весовой коэффициент. В отличие от полносвязного, в сверточном (convolution) слое нейрон соединен лишь с ограниченным количеством нейронов предыдущего уровня, то есть сверточный слой аналогичен применению операции свертки, где используется лишь матрица весов небольшого размера (ядро свертки).

Субдискретизирующий слой выполняет уменьшение размерности (pooling). Это можно делать различными способами, но зачастую используется метод выбора максимального элемента - слой признаков разделяется на ячейки, из которых выбирается максимальное значение и передается в следующий слой.

Dropout слой представляет собой метод регуляризации в нейронных сетях, обучение в которых обычно производят стохастическим градиентным спуском. Dropout регуляция заключается в изменении структуры сети: каждый нейрон отключается с некоторой вероятностью p на i-ом этапе обучения. По такой прореженной сети производится обучение, а для оставшихся весов делается градиентный шаг.

Ансамбль нейронных сетей

Точность идентификации объектов может быть повышена с использованием АНС [1-3], то есть набора НС, принимающего решения путем усреднения результатов работы отдельных НС. Для этого необходимо реализовать вариативность НС в ансамбле. Для чего могут использоваться следующие подходы или их сочетания:

- использование различных частей обучающей выборки;
- случайная инициализация весов НС;
- вариации архитектур НС в ансамбле (добавление скрытых слоев, добавление/удаление нейронов скрытого слоя).

Выходное значение АНС формируется как взвешенная сумма выходов отдельных НС. Данный подход проиллюстрирован на рисунке 1.

Случай с одним выходным нейроном вычисляется по формуле:

$$y = \sum_{i=1}^{n} y_i \cdot w_i \quad (1)$$

где n - число моделей; y_i - выход i-й НС; w_i - вес i-й НС, который формируется по формуле:

$$w_i = \frac{A_i}{\sum_{j=1}^{n} A_j}, \quad (2)$$

где A_i - выбранная мера ошибки, рассчитанная для i-й НС;

n - число моделей.

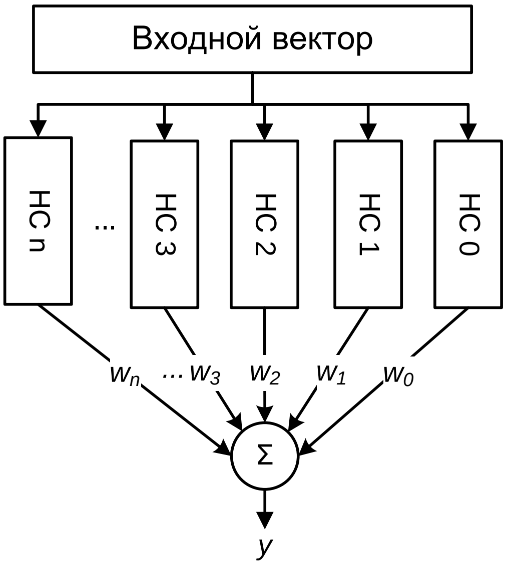


Рис. 1. Взвешенный АНС

Методика на основе ансамблей моделей для идентификации объектов дистанционного зондирования Земли

Предлагаемая методика состоит из следующих шагов:

Шаг 1. Постановка задачи: определение типа исходных данных, определение идентифицируемых объектов, определение меры оценки качества моделей.

Шаг 2. Формирование обучающего набора на исходных данных: сбор данных, предобработка, разметка выходного набора.

Шаг 3. Поиск дополнительных признаков для решения поставленной задачи.

Шаг 4. Расширение обучающего набора дополнительными признаками.

Шаг 5. Разделение обучающего набора на тренировочный и тестовый.

Шаг 6. Определение архитектуры модели на основе исходных данных.

Шаг 7. Определение диапазонов гиперпараметров выбранной архитектуры.

Шаг 8. Определение оптимальных гиперпараметров модели методами сеточого или случайного поиска [6] с использованием k-кратной перекрёстной проверки на тренировочном наборе.

Шаг 9. Если модель удовлетворяет мере оценки качества на тестовом наборе, то задача решена, иначе необходимо расширить набор данных и перейти на шаг 4.

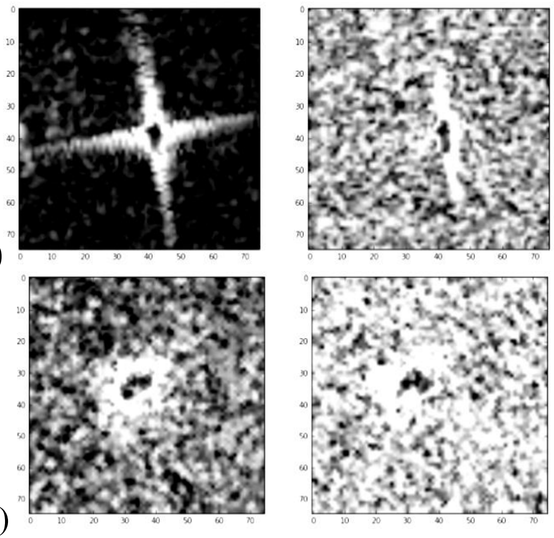
Проведение эксперимента

Набор экспериментальных данных представляет собой изображения, полученные с помощью радара с синтезированной апертурой (РСА), который позволяет получать радиолокационные изображения земной поверхности и находящихся на ней объектов независимо от метеорологических условий и уровня естественной освещенности местности с детальностью, сравнимой с аэрофотоснимками.

Данные [4]:

– снимки в двух поляризационных режимах: горизонтальный/горизонтальный, горизонтальный/вертикальный;
– каждый снимок содержит один объект: корабль или айсберг (см. рисунок Рис. 2);
– угол падения;
– тренировочный набор: 1604 снимков, 75x75;
– тестовый набор: 8424 снимка, 75x75.

Необходимо решить задачу бинарной классификации айсберг/корабль.

а)
б)

Рис. 2. Пример изображений: а) корабль, б) айсберг

Меры оценки моделей

Эффективность в задаче классификации может оцениваться с использованием матрицы неточностей (confusion matrix). С ее помощью можно рассчитать такие метрики, как точность (*Precision*), полнота (*Recall*) и *F*-мера. Базовой мерой выступает *accuracy* - доля правильных ответов модели.

Матрица неточностей - это матрица размера N на N, где N - это количество классов. Столбцы этой матрицы резервируются за фактическими значениями индексов классов, а строки – за решениями классификатора. Когда классифицируется объект из тестовой выборки, то увеличивается значение, стоящее на пересечении строки класса, который вернул классификатор, и столбца класса, к которому действительно относится объект. Точность равна отношению соответствующего диагонального элемента матрицы и суммы всей строки класса. Полнота – отношению диагонального элемента матрицы и суммы всего столбца класса.

F-мера представляет собой гармоническое среднее между точностью и полнотой.

Для задачи бинарной классификации, при выводе моделью вероятности класса, может быть использована логарифмическая функция потерь:

$$logloss = -\frac{1}{l}\sum_{i=1}^{l}(y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)) \quad (3)$$

где $\hat{y}_i$ - это ответ алгоритма на i-ом объекте, y - истинная метка класса на i-ом объекте, l - размер выборки.

Минимизацию *logloss* можно представить как задачу максимизации *accuracy* путем штрафа за неверные предсказания. Однако необходимо отметить, что *logloss* крайне сильно штрафует за уверенность классификатора в неверном ответе. Следовательно, ошибка на одном объекте может дать существенное ухудшение общей ошибки на выборке. Такие объекты часто бывают выбросами, которые нужно не забывать фильтровать или рассматривать отдельно.

Результаты

Исходя из параметров исходных данных, была предложена базовая СНС:

Сверточный слой, 64 ядра 3х3;
Pooling 2x2;
Dropout с вероятностью 0.3;
Сверточный слой, 128 ядер 3х3;
Pooling 2x2;
Dropout с вероятностью 0.3;
Сверточный слой, 128 ядер 3х3;
Pooling 2x2;
Dropout с вероятностью 0.3;
Сверточный слой, 64 ядер 3х3;
Pooling 2x2;
Dropout с вероятностью 0.3;
Полносвязный слой 1024 нейрона;
Dropout с вероятностью 0.1;
Полносвязный слой 512 нейрона;
Dropout с вероятностью 0.1;
Выходной слой softmax.

СНС принимает на вход псевдоизображение, в котором первый канал изображения представлен ГГ каналом исходных данных, второй канал изображения представлен ГВ каналом, третий канал изображения представлен их композитом. Обучение СНС выполнялось по стохастическому градиентному алгоритму Adam [8]. Перекрестная проверка выполнялась для k = 5. На основе базовой СНС формировался АНС из пяти моделей.

Результат оценки данных моделей представлен в таблице 1. Из таблицы видно, что Ансамбль СНС моделей показывает наименьшую *logloss* ошибку на тестовом наборе данных.

Таблица 1. Результат оценки моделей

Модель	*Accuracy, тренировочный набор*	*Logloss, тренировочный набор*	*Accuracy, тестовый набор*	*Logloss, тестовый набор*
СНС	0.922	0.203	0.917	0.223
Ансамбль	0.936	0.155	0.932	0.166

Заключение

Предложена методика для идентификации объектов дистанционного зондирования Земли. Методика включает: подготовку данных, определение признаков объектов, формирование архитектуры модели идентификации и поиск гиперпараметров модели. Предложены мо-

дели для анализа изображений ДЗЗ на основе ансамбля СНС.

Методика была применена при проведении эксперимента с использованием данных ДЗЗ. Оценка предложенных моделей показала, что одним из эффективных способов увеличения точности в задачах машинного обучения является формирование ансамбля моделей.

References

1. Parikh, D., Polikar, D. An ensemble-based incremental learning approach to data fusion (2007). IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics. Vol. 37. Iss. 2. P. 437 450.
2. Marushko, Y. Using Ensembles of Neural Networks with Different Scales of Input Data for the Analysis of Telemetry Data (2013). Proc. of the XV Intern. PhD Workshop OWD 2013 (Wisla, 19-22 Oct. 2013). Gliwice: Silesian University of Technology. P. 386-391.
3. Kourentzes, N. Neural network ensemble operators for time series forecasting (2014). Expert Systems with Applications. July 2014. Vol. 41, Iss. 9. P. 4235-4244.
4. Statoil/C-CORE Iceberg Classifier Challenge. Data [Online]. Available: https://www.kaggle.com/c/statoil-iceberg-classifier-challenge/data.
5. Large Scale Visual Recognition Challenge [Online]. Available: http://image-net.org/challenges/LSVRC/2016/results.
6. Bengio, Y., Bergstra, J. Random Search for Hyper-Parameter Optimization (2012). J. Machine Learning Research. 13. P. 281-305.
7. Goodfellow, Ia., Bengio, Y., Courville, A. Deep Learning (2016). MIT Press. 781 p.
8. Diederik, P., Kingma. A., Ba, J. A Method for Stochastic Optimization (2015). 3rd International Conference for Learning Representations, San Diego.

RESUME

Y. Marushko, A. Doudkin

Using an ensemble of convolutional neural networks to process remote sensing data

A method for identifying objects of remote sensing of the Earth is proposed, the method includes: determining the type of source data, determining identifiable objects, determining a measure for assessing the quality of models, forming a training set on the source data, determining features of objects, forming an identification model architecture, determining ranges of hyper-parameters of the selected architecture hyper-parameters of the model. It is proposed to perform the determination of the optimal hyperparameters of the model using the grid or random search methods using the k-fold cross-check on the training set.

Models for the analysis of images of remote sensing of the Earth on the basis of an ensemble of convolutional neural networks are proposed. The architecture of convolutional neural networks is an alternation of convolutional layers, subsampled layers, and fully connected layers.

The accuracy of identifying objects can be improved using an ensemble of neural networks, that is, a set of models that makes decisions by aggregating the results of the work of individual networks. In this paper, the output value of an ensemble of neural networks is formed as a weighted sum of the outputs of individual models.

The technique was applied during the experiment using data from remote sensing of the Earth. The data set is an image obtained using synthetic aperture radar.

The basic convolutional neural network contains four convolutional layers and three full-stranded. The output layer is softmax. The network receives as input a pseudo-image in which the first channel of the image is represented by the horizontal / horizontal channel of polarization of the source data, the second channel is represented by the images by the horizontal-vertical channel of polarization, the third channel of the image is represented by their composite as normalized sum of channels.

Network training was performed using the Adam stochastic gradient algorithm. Based on the basic convolutional network, an ensemble of five models was developed that were trained on different parts of the initial data in the process of cross-validation.

Evaluation of the proposed models showed that one of the effective ways to increase accuracy in machine learning tasks is to form an ensemble of models.

Надійшла до редакції 15.11.2018